

Wikipedia como vocabulario controlado: ¿está superado el control de autoridades tradicional?

Por Jesús Jiménez-Pelayo

Resumen: Wikipedia, la enciclopedia libre, es el primer proyecto específicamente nacido desde y para la Web que ha desarrollado un sistema de control de autoridades para el acceso a su información. En este trabajo se analizan los elementos, procedimientos y principios que lo constituyen, y se traza una analogía crítica entre éstos y los que conforman el control de autoridades tradicional aplicado a los catálogos bibliográficos. A partir de la comparación crítica de ambos modelos se plantea hasta qué punto el control de autoridades bibliotecario, esclerotizado por el peso de la tradición y por su nula adaptación a la tecnología, ha sido superado por proyectos como el de Wikipedia, que parten de una filosofía de la flexibilidad y el sentido común, y donde las normas se deciden por y para el usuario. El enorme potencial y alcance del modelo de autoridades de Wikipedia hacen que se perfila como el gran candidato a convertirse en el sistema de acceso normalizado a la web semántica.

Palabras claves: Wikipedia, Vocabularios controlados, Control de autoridades, Web semántica, Normalización de información.

Title: *Wikipedia as controlled vocabulary: has the traditional authority control been surpassed?*

Abstract: Wikipedia, the free encyclopaedia, is the first project to have been born specifically from and for the web and which has developed an authority control system for access to their information. Here, the different elements, procedures and principles which make up the Wikipedia authority control system is analysed, and a critical analogy is traced between them and those that meet the traditional authority control applied to bibliographical catalogues. From the critical comparison of both models, one wonders up to what point the authority control, constrained by the weight of tradition and by its hopeless adaptation to technology, has been surpassed by developments such as Wikipedia which are based on a philosophy of flexibility and common sense and where the rules are decided by and for the user. The enormous potential and reach of the Wikipedia authority model make it a true contender to become the normal system of access to the Semantic Web.

Keywords: Wikipedia, Controlled vocabulary, Authority control, Semantic web, Information normalization.

Jiménez-Pelayo, Jesús. "Wikipedia como vocabulario controlado: ¿está superado el control de autoridades tradicional?". *El profesional de la información*, 2009, marzo-abril, v. 18, n. 2, pp. 188-201.

DOI: 10.3145/epi.2009.mar.09



Jesús Jiménez Pelayo, licenciado en documentación y diplomado en biblioteconomía y documentación por la Universidad de Granada, es coordinador de la Biblioteca Virtual de Andalucía. Autor de monografías y estudios sobre normalización de información, entre los cuales destaca "El catálogo de autoridades, creación y gestión en unidades documentales". Colabora habitualmente como docente en cursos, y ha sido miembro del Comité Permanente de Clasificación e Indización de la Federación Internacional de Bibliotecas e Instituciones Bibliotecarias (IFLA), representando a España.

1. Introducción

Wikipedia, la enciclopedia libre, es el primer proyecto específicamente nacido desde y para la Web que ha elaborado un vocabulario controlado para el acceso a su contenido¹. Fundada en 2001 por los estadounidenses **Jimmy Wales** y **Larry Sanger**, se autodefine como libre y políglota, basada en la colaboración de sus contribuyentes por medio de la *tecnología wiki*. Su objetivo, en palabras de su co-fundador **Jimmy Wales**, es "crear y distribuir una enciclopedia libre, de la más alta calidad posible, a cada persona del planeta y en su idioma, para lograr un mundo en el que todos tengamos libre acceso a la suma de todo el saber

de la humanidad"². Pese a lo ambicioso de su meta, lo cierto es que se va consolidando con paso firme. Actualmente presenta ediciones en más de 253 idiomas, la primera de las cuales en volumen es la inglesa, con cerca de 3 millones de artículos, seguida de la alemana y la francesa con cerca de 1 millón; la española se encuentra en el noveno lugar con más de 400.000 artículos.

La popularidad cada vez mayor de Wikipedia como fuente de información, su continuo crecimiento y su mejora constante hacen que resulte interesante plantear y conocer cuáles son las soluciones propuestas en el terreno de la normalización terminológica por un gigante

Artículo recibido el 28-09-08

Aceptación definitiva: 16-12-08

de la Web. El interés es doble si tenemos en cuenta que dicha normalización no proviene de un entorno bibliotecario-documentalista, sino de un sistema colaborativo que se autogestiona, en el que no hay normas impuestas sino consensuadas, y donde son los propios usuarios los que están escribiendo las reglas que deben regirlos día a día.

“Wikipedia es el primer proyecto específicamente nacido desde y para la Web, con un vocabulario controlado para el acceso a su contenido”

A continuación se analizan los elementos, procedimientos y principios que constituyen el sistema de control del vocabulario o control de autoridades de *Wikipedia*, y se traza una analogía crítica entre éstos y los que conforman el sistema tradicional aplicado a los catálogos bibliográficos. A partir de la comparación de ambos modelos de acceso a la información nos planteamos hasta qué punto el control de autoridades tradicional ha quedado desfasado frente a iniciativas como *Wikipedia*, que parten de una filosofía de la flexibilidad y el sentido común, y donde las normas se deciden por y para el usuario.

2. Control de autoridades en Wikipedia

Con *Wikipedia* se ha elaborado un sistema propio de control de autoridades que permite a sus usuarios acceder de forma cómoda, fiable y organizada al contenido de sus artículos. Como cualquier sistema de este tipo, por una parte está formado por una serie de elementos y procedimientos que lo hacen posible técnicamente y por otra se rige por unos principios básicos, que son los que en realidad definen sus pautas de normalización.

No puede decirse que los elementos y procedimientos usados para controlar autoridades sean novedosos ya que, con distintos nombres y con algunas diferencias en su método, son muy parecidos a los que se emplean en catálogos bibliográficos. Sin embargo, sí que son novedosos los principios que rigen la determinación, construcción y organización de las autoridades, ya que adoptan soluciones originales que se apartan por completo de la normativa bibliotecaria. En la tabla 1 se ofrece una relación de los distintos elementos y procedimientos que componen el control de autoridades de *Wikipedia* y de los catálogos bibliográficos respectivamente; en él se equiparan los términos homólogos de uno y otro modelo.

CONTROL DE AUTORIDADES Elementos y procedimientos comunes	
Catálogo bibliográfico	Wikipedia
Autoridades	Entradas normalizadas
Autoridades paralelas	Entradas paralelas
Notas de información	<i>Hatnotes</i>
Referencias de véase	Redirecciones
Referencias de véase múltiples	Páginas de desambiguación
Referencias de véase además	Véase también
Relaciones jerárquicas	Categorías

Tabla 1

2.1. Entradas normalizadas: las autoridades de Wikipedia

En *Wikipedia* se normalizan los títulos de los artículos (entradas), lo que hace que esos elementos funcionen como verdaderas autoridades de su sistema de información. Las entradas pueden ser nombres de persona, nombres de entidad, lugares geográficos, títulos de obras, o materias; es decir, todos aquellos elementos de información que tradicionalmente se consideran susceptibles de ser sometidos a un proceso de normalización o control de autoridades.

La entrada normalizada o autoridad, figura de manera destacada en la cabecera de cada artículo (figura 1). Para su elección y redacción se siguen una serie de convenciones que aparecen recogidas en manuales como: *Wikipedia: naming conventions*³.

A continuación abordaremos el estudio de las principales directrices para la normalización de las entradas, a cada una de las cuales contraponemos un principio básico que rige la normativa para autoridades de los catálogos bibliográficos. De esa forma el lector podrá hacerse una idea más exacta de la novedad que representa *Wikipedia* respecto a los productos tradicionales.

a) Normalización consensuada por los usuarios vs. normalización impuesta por los bibliotecarios

En la normalización de autoridades tradicional el bibliotecario es la figura clave. Es el agente encargado de “forzar” los datos de los documentos, de someterlos a un proceso de control –denominado control de autoridades– con el fin de establecer puntos de acceso uniformes que puedan funcionar siempre como claves seguras de búsqueda (autoridades). Los bibliotecarios elaboran herramientas de búsqueda tales como los catálogos bibliográficos, mediante la imposición de unos estándares (códigos de catalogación y vocabularios controlados) creados por ellos mismos. En su proceso de confección el usuario no interviene en absoluto, sólo los usa.

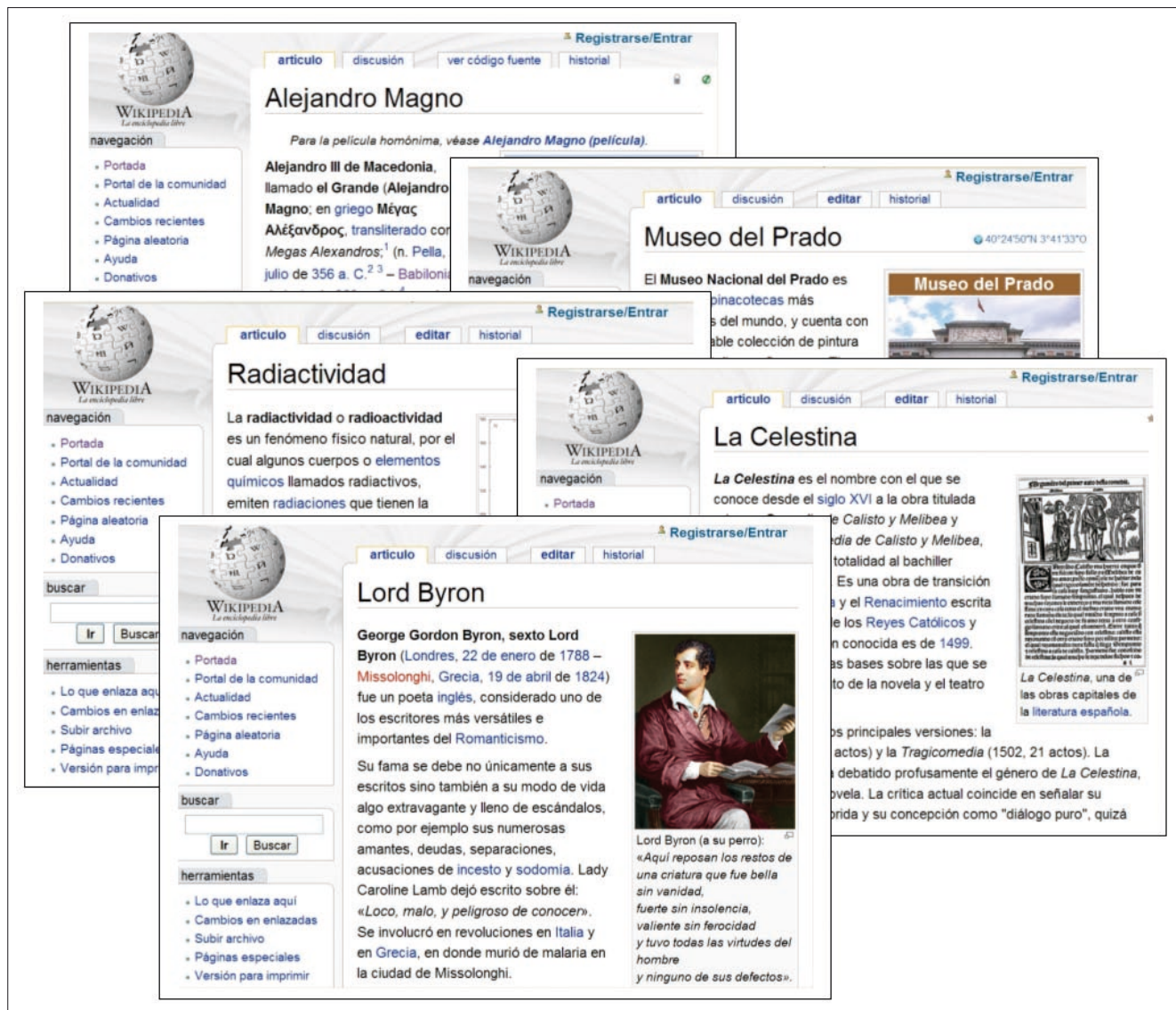


Figura 1

“En Wikipedia se respeta la intervención del usuario en el proceso de control de información”

Un modelo completamente distinto es el que ofrece Wikipedia. Se prescinde de la figura del bibliotecario⁴ y son los propios usuarios de la enciclopedia los que “toman las riendas” en la organización y normalización de la información. Esto incluye la adopción de forma consensuada de una serie de convenciones para determinar y redactar el nombre de sus entradas o autoridades. Cada edición (idioma) de Wikipedia adopta sus propias políticas, aunque algunas son comunes. Una vez que la comunidad logra el consenso sobre la aplicación de una norma, todos los editores están obligados a respetarla. En las directrices de Wikipedia para la normalización de sus términos subyace un respeto

absoluto por la intervención del usuario en el proceso de control de información, por que sea éste —el usuario como colectivo— y no otro agente, quien elija de manera consensuada y según sus propios usos lingüísticos y alfabéticos la forma autorizada de un nombre, un título o una materia.

Esta idea constituye uno de los principios básicos y aparece formulada del siguiente modo en sus *Naming conventions*: “Los nombres de los artículos de Wikipedia deberían estar pensados más para lectores que para editores, y más para un público general que para especialistas”⁵.

b) Flexibilidad vs. rigidez

Wikipedia es consciente de su rápido crecimiento y continuo cambio, por lo que en sus *Naming conventions* se recalca el hecho de que proporciona convenciones para establecer autoridades y “no normas grabadas en piedra”³. Las convenciones se revisarán constantemente y, en caso de que con el tiempo se queden obsoletas o

inapropiadas, se modificarán o se cambiarán por otras, buscando siempre que el usuario acceda a la información de manera cómoda, sencilla y rápida. Se advierte además que “a veces deberá primar el sentido común o la visión práctica sobre las reglas establecidas”. No es precisamente este afán de puesta al día y de adaptación a las necesidades de la gente lo que caracteriza la normativa de autoridades -que sí parece estar grabada en piedra-, aplicada a los catálogos bibliográficos. La mayor prueba del espíritu de flexibilidad de *Wikipedia* la encontramos en su declaración de principios: “*Wikipedia* no hace la realidad, es la realidad la que hace y transforma *Wikipedia*; y la realidad son sus usuarios”⁶.

“Wikipedia proporciona convenciones para establecer autoridades y no normas grabadas en piedra”

c) Simplicidad vs. complejidad

Las convenciones para autoridades son pocas y fáciles de interpretar y aplicar. Su manual para la redacción de autoridades de la edición española incluye en total unas 15 reglas: 5 generales para todo tipo de autoridad, y 10 específicas para casos concretos como nombres de persona, títulos, organismos, etc. La versión inglesa es más amplia, cuenta con unas 50 directrices, aunque en su mayoría se trata más de ejemplificaciones de aplicación de los principios generales a casos particulares que de reglas propiamente dichas.

En cualquier caso, resultan sorprendentes estas cifras si las comparamos con el número de reglas que componen los códigos de catalogación. En las *Reglas de catalogación* españolas, se recogen más de 200 especificaciones normativas distintas referidas a personas, entidades y títulos⁷. Si a ello le sumamos las de materias y nombres geográficos que establecen las distintas listas de encabezamientos de materias y tesauros aplicados a catálogos, podemos cifrar en torno a 300 las reglas que un bibliotecario maneja para normalizar puntos de acceso⁸. Esa diferencia numérica tiene su explicación en que *Wikipedia* da directrices muy generales y se basa en todo momento en la iniciativa libre, el sentido común y la visión práctica de los wikipedistas a la hora de asignar autoridades. Además opta por principios tajantes como el de redactar las autoridades en la forma más conocida en el idioma del usuario, o usar siempre el orden directo para sus entradas, a lo que hay que añadir su tendencia a construir las autoridades de la forma más simple posible, desprovistas de elementos añadidos al nombre (tales como términos de tratamien-

to, fechas, etc.). Estos tres principios, de los que hablaremos más adelante, son totalmente innovadores y rompen con la gran mayoría de las reglas que tradicionalmente se han venido considerando imprescindibles para un tratamiento analítico y particularizado de las autoridades, que son precisamente las que conforman la mayor parte del articulado normativo de cualquier código de catalogación.

“El sentido común, la espontaneidad y la visión práctica del usuario definen este modelo novedoso”

d) Naturalidad vs. artificialidad

Wikipedia aboga también por la naturalidad en el acceso a su información. Para ello no sólo trabaja en que sus artículos estén bien estructurados, presentados y vinculados, sino en que su contenido sea accesible de modo intuitivo. Las autoridades juegan un papel clave en ese acceso. Por eso se pone especial atención a que cada autoridad sea “una representación lo más cercana posible a lo que lo que la gran mayoría de los usuarios espera encontrar como autoridad cada vez que hace una búsqueda”⁵. De esta forma procura evitar al máximo la distorsión entre lo que se busca y lo que se encuentra.

Se tiene en cuenta cómo los usuarios enuncian mentalmente sus búsquedas e intenta que sus autoridades se aproximen a esos enunciados, eligiendo de entre todas las variantes posibles de designación aquellas que resultan más fácilmente reconocibles o más naturales para el usuario. Se aparta así de nuevo de los catálogos bibliográficos, cuyos estándares crean en ocasiones situaciones de gran artificialidad, al adoptar autoridades que bien pueden ser calificadas de difíciles de prever, insospechables o rebuscadas. En la tabla 2 se recogen algunos ejemplos de esta contraposición naturalidad-artificialidad entre *Wikipedia* (a la izquierda) y catálogos bibliográficos (a la derecha)⁹.

A la vista de estos ejemplos estaremos de acuerdo en que las probabilidades de que un usuario interesado en obtener información sobre el pensador **Montesquieu** enunciara mentalmente su búsqueda como *Montesquieu* (autoridad en *Wikipedia*) serían muy altas, mientras que resultaría bastante improbable que su enunciado mental coincidiera con *Montesquieu, Charles de Secondat, baron de* (autoridad en los catálogos bibliográficos). Lo mismo podemos decir de los restantes casos, donde destacan autoridades tan extravagantes como *Savoia, Luigi Amedeo di, duca degli Abruzzi*, para el navegante y explorador **Luis Amadeo de Sabo-**

Wikipedia	Catálogos bibliográficos
Montesquieu	Montesquieu, Charles de Secondat, baron de
Luis Amadeo de Saboya	Savoia, Luigi Amedeo di, duca degli Abruzzi
Sofía de Grecia	Sofía, Reina consorte de Juan Carlos I, Rey de España
Elogio de la locura	Erasmus, Desiderius. <i>Morias Enkomion</i> . Español
Pedro Arrupe	Jesuitas. <i>Prepósito General</i> (1965-1983: Pedro Arrupe)
Desierto del Sahara	Sahara (Desierto)
Concierto para clarinete (Mozart)	Mozart, Wolfgang Amadeus. <i>Conciertos, clarinete, orquesta</i> , K. 622, la mayor
Boecio	Boecio, Anicio Manlio Torcuato Severino
Guerra civil española	España-Historia-Guerra civil, 1936-1939

Tabla 2

ya; *Erasmus, Desiderius. Morias Enkomion. Español*, para la obra de **Erasmus** *Elogio de la locura*; o *Jesuitas. Prepósito General* (1965-1983: **Pedro Arrupe**), para el jesuita **Pedro Arrupe**.

La naturalidad en las autoridades se consigue mediante dos principios básicos que veremos a continuación: 1) El uso del idioma del usuario; y 2) La redacción en orden directo.

e) Idioma del usuario vs. idioma original

Precisamente porque los usuarios son los que marcan la realidad de *Wikipedia*, el principio general que rige la determinación de sus autoridades es bien sencillo, claro y de sentido común. La elección de la autoridad recaerá en la forma más conocida para designar un nombre, título o materia en el idioma y alfabeto del usuario. La única excepción a esto se encuentra obviamente en los nombres o títulos para los que no exista una forma acuñada o afianzada en la lengua en cuestión. La diferencia con respecto a la normativa de los catálogos bibliográficos es importante, ya que en ésta la autoridad recae igualmente sobre la forma del nombre más comúnmente usada, pero redactada en su lengua original.

El principio de *Wikipedia* es bastante sólido, se acomoda perfectamente al usuario y crea una única y

lógica excepción. El de los catálogos bibliográficos tiene el grave defecto de que a esa regla general incorpora numerosas excepciones, como las previstas para tipologías concretas como materias, lugares, nombres de santos, beatos, papas, emperadores, reyes, príncipes reinantes, clásicos griegos y latinos, entidades internacionales, títulos de obras anónimas y de obras musicales, etc., para las cuales se prefiere el uso de la lengua del usuario. Esta dualidad hace que al final la percepción de una normativa así sea la de algo poco definido, algo que se queda entre dos aguas, sin llegar a respetar plenamente la forma original del nombre, ni tampoco la forma más conocida por los usuarios. En la tabla 3 se ofrecen algunos ejemplos comparativos que ilustran ambos sistemas.

Wikipedia	Catálogos bibliográficos
Julio Verne	Verne, Jules
Museo Británico	British Museum
Orquesta Filarmónica de Berlín	Berliner Philharmoniker
Los viajes de Gulliver	Swift, Jonathan. <i>Gulliver's travel</i> . Español
El sexto sentido	The sixth sense

Tabla 3

f) Orden directo vs. orden directo/inverso

Las autoridades de *Wikipedia* van siempre redactadas en orden directo, ya se refieran a nombres de personas, instituciones, lugares geográficos, títulos o materias. Es un principio radical que busca la naturalidad en la expresión de cada autoridad, al mismo tiempo que no genera dudas, conflictos o excepciones en su interpretación. Esto choca frontalmente con los catálogos bibliográficos, donde la convivencia de autoridades en orden directo e inverso es una práctica habitual, propiciada por una normativa que recoge reglas específicas, a veces complejas y difíciles de interpretar, relativas a determinar cuál debe ser el elemento inicial de entrada de una autoridad formada por varios elementos.

Los bibliotecarios han venido heredando esas reglas durante generaciones y las aplican al catálogo de manera escrupulosa. Desde una perspectiva actual, esta dualidad orden directo/orden inverso no resiste una revisión crítica, por mínima que sea. Hoy resulta difícil encontrar argumentos sólidos que sostengan por qué en catálogos bibliográficos la autoridad para **Federico García Lorca** es *García Lorca, Federico* (orden inverso), mientras que para **Gonzalo de Berceo** es *Gonzalo de Berceo* (orden directo). Por muchas razones técnicas que buscáramos, la única explicación posible es la de tipo histórico, es decir, el enorme peso que la tradición

“Wikipedia recrea el modo en que los usuarios enuncian mentalmente sus búsquedas e intenta que sus autoridades se aproximen a esos enunciados”

tiene en la normativa que rige los catálogos bibliográficos. Esta dicotomía en el orden de redacción de las autoridades tuvo su justificación, pero hay que retrotraerse mucho para encontrarla, ya que proviene de los antiguos catálogos manuales y de la costumbre de alfabeticar sus asientos para mayor comodidad del usuario, utilizando como elemento inicial de entrada la parte del nombre que presuntamente es más susceptible de ser buscada. Desde que esta práctica fue enunciada por **Panizzi**¹⁰ en 1841, se ha venido arrastrando en el tiempo e incorporando en todos los códigos de catalogación. Pero ni el usuario de 1841 es el usuario del siglo XXI, ni la tecnología del fichero de gaveta y cartulina tiene nada que ver con la tecnología web. Por eso frente a la apolillada normativa de los catálogos bibliográficos, el principio de *Wikipedia* contrasta por su gran espontaneidad, frescura y sobre todo, criterio único.

La redacción de los nombres de persona no puede ser más sencilla, pues no hay que tener en cuenta si el nombre consta de uno o varios apellidos, presenta partículas (preposiciones, artículos), sobrenombres, etc., para decidir su orden de entrada. En la tabla 4 mostramos unos ejemplos, a los que contraponemos las soluciones adoptadas en los catálogos bibliográficos.

Wikipedia	Catálogos bibliográficos
Federico García Lorca	García Lorca, Federico
Jean de La Fontaine	La Fontaine, Jean de
Eugenio d'Ors	Ors, Eugenio d'
Edgar Allan Poe	Poe, Edgar Allan
Leonardo da Vinci	Leonardo da Vinci

Tabla 4

En los catálogos bibliográficos las instituciones de carácter administrativo reciben un tratamiento controvertido y farragoso ya que, dependiendo de su finalidad, unas veces se redactan en orden directo, y otras en orden inverso introducidas por el lugar. Este tratamiento queda descartado en *Wikipedia*, donde todas las entidades, privadas o públicas, administrativas o no, autónomas o dependientes, son redactadas en directo. Algunos ejemplos se muestran en la tabla 5.

El tratamiento de los títulos con entrada directa produce resultados que merecen algún comentario (tabla 6).

A diferencia de los catálogos bibliográficos, las autoridades de título de *Wikipedia* constan solamente del título normalizado, no van precedidas del nombre del autor ni de ningún otro elemento de información. En realidad es innecesario ya que el autor está relacionado con todos los títulos de sus obras y al revés a través de

Wikipedia	Catálogos bibliográficos
Ayuntamiento de Madrid	Madrid. Ayuntamiento
Ministerio de Cultura de España	España. Ministerio de Cultura
Parlamento de La Rioja	La Rioja (Comunidad autónoma). Parlamento
Congreso de los Estados Unidos	Estados Unidos. Congress
Facultad de Medicina (Universidad de Sevilla)	Universidad de Sevilla. Facultad de Medicina
Consejería de Innovación, Ciencia y Empresa de Andalucía	Andalucía. Consejería de Innovación, Ciencia y Empresa

Tabla 5

Wikipedia	Catálogos bibliográficos
El mercader de Venecia	Shakespeare, William. Merchant of Venice. Español
Tintín en el país del oro negro	Hergé. Tintin au Pays de l'Or Noir. Español
Ley de Extranjería de España	España. Ley de extranjería, 2000
Libro de Job	Biblia. A.T. Libro de Job
Estatuto de autonomía de Andalucía	Andalucía. Estatuto de autonomía, 2007
Sinfonía nº 3 (Beethoven)	Beethoven, Ludwig van. Sinfonías, n. 3, op. 55, mi bemol mayor

Tabla 6

los vínculos que aparecen en el texto del artículo, lo que hace que la información esté conectada de manera permanente, sin que esa conexión se muestre de manera explícita al usuario. Sólo en los casos en que hay homonimia y se produce ambigüedad, el título se acompaña de un calificador para su diferenciación. Hay que destacar también la sencillez de construcción de los títulos de obras musicales, de tratamiento tan extremadamente complejo en los catálogos bibliográficos.

Nombres geográficos y materias no son una excepción a la regla de la entrada directa (tabla 7).

Wikipedia	Catálogos bibliográficos
Historia de Grecia	Grecia-Historia
Filosofía del Derecho	Derecho-Filosofía
Revolución rusa de 1917	Rusia-Historia-1917-1921
Colonización española de América	América Latina-Colonización-España
Mar Caspio	Caspio (Mar)
Pico del Veleta	Veleta (Pico, España)

Tabla 7

En el caso de las últimas, *Wikipedia* se aparta de los tradicionales encabezamientos de materia constituidos por enunciados del tipo “Encabezamiento-Subencabezamiento” para acercarse a la filosofía de los descriptores: como tales hay que considerar no sólo los términos temáticos que son entrada normalizada de sus artículos, sino también las categorías que agrupan estas entradas (véase apartado 2.7. Categorías).

Cabe preguntarse hasta qué punto el crecimiento continuo y la necesidad futura de normalizar autoridades cada vez más específicas hará que este principio de entrada en orden directo resulte ineficaz, o incluso que se agote en sí mismo por exceso de nombres homónimos. *Wikipedia* cuenta como baza con las enormes posibilidades de los enlaces para conectar autoridades a través del cuerpo informativo de cada artículo, como ocurre en el caso de los autores y los títulos de sus obras. En su contra, sin embargo, tiene el hecho de que una saturación de homonimias exige un mayor uso de calificadores y más amplias páginas de desambiguación para dar cabida a todas las posibles diferenciaciones (véase epígrafe 2.5.). El tiempo dirá.

“Los catálogos bibliográficos están poco pensados para el usuario”

g) No adición vs. adición

Wikipedia opta por construir autoridades desprovistas de términos añadidos al nombre que especifiquen rango, tratamiento, cargo, condición social, etc., tales como santo, beato, fray, sor, rey, papa, obispo, duque, marqués, infanta, sir, etc. A diferencia de los catálogos bibliográficos, donde son frecuentes estas adiciones, en

	Wikipedia	Catálogos bibliográficos
	Marco Aurelio	Marco Aurelio, Emperador de Roma
	Carolina de Mónaco	Grimaldi, Caroline, princesse de Monaco
	Teresa de Jesús	Teresa de Jesús, Santa
	Benedicto XVI	Benedicto XVI, Papa
	Luis XVI de Francia	Luis XVI, Rey de Francia
(excepción)	Marqués de Sade	Sade, Donatien Alphonse François, marquis de
(excepción)	Arcipreste de Hita	Ruiz, Juan (s. XIV)

Tabla 8

Wikipedia sólo se utilizan en los casos excepcionales en que la persona sea conocida por el nombre que incluye el término (tabla 8).

h) Calificación vs. subordinación

En las autoridades de *Wikipedia* no existe subordinación de elementos como ocurre en los catálogos bibliográficos, donde son típicas las construcciones “Autor. Título”, “Jurisdicción. Nombre de entidad”, “Entidad principal. Entidad dependiente”. Esto da lugar a mayor número de nombres homónimos, cuyo significado *Wikipedia* diferencia con el uso de calificadores, que son mucho más frecuentes y diversos que en los catálogos bibliográficos. Van siempre entre paréntesis después de la autoridad y pueden ser desde un nombre de persona o entidad, hasta términos que designan profesión, parentesco, género literario, tipología documental, disciplinas, lugares, etc. No se utilizan nunca fechas de nacimiento y/o muerte. Es más frecuente encontrarlos en títulos de obras y materias (tabla 9).

Alexandre Dumas	(hijo)
David Copperfield	(novela)
Concierto para piano nº 1	(Chopin)
Dracula	(película de 1958)
Granada	(canción)
Triángulo	(instrumento musical)
Facultad de Medicina	(Universidad de Granada)

Tabla 9

i) Unicidad vs. entrada múltiple

La normativa de autoridades para catálogos bibliográficos estipula la creación de varias entradas de autoridad para una misma persona o entidad en casos tales como: seudónimos no predominantes, distintos nombres según el tipo de obra o función realizada, cargos políticos y religiosos, y entidades que cambian de nombre a lo largo del tiempo. Este procedimiento de entrada múltiple que va en contra de la unicidad del propio concepto de autoridad, no existe en *Wikipedia* donde se crea siempre, sin excepción, una única entrada para cada persona o entidad. Un ejemplo de este diferente tratamiento lo vemos en la tabla 10.

j) Singular vs. plural

Wikipedia normaliza sus entradas de materia siempre en singular, a diferencia de los catálogos bibliográficos.

Wikipedia	Catálogos bibliográficos
Alfonso X el Sabio	Alfonso X, Rey de Castilla
	Castilla (Reino). Rey (1252-1284: Alfonso X)

Tabla 10

cos, donde las listas de encabezamientos y tesauros que lo rigen obligan al uso del plural para una amplia mayoría de términos, concretamente para los que designan realidades sustantivas contables. Así, frente a materias como, por ejemplo, *automóviles*, *enzimas*, *perfumes*..., *Wikipedia* prefiere *automóvil*, *enzima* y *perfume*...

2.2. Entradas paralelas

Son otras formas lingüísticas en las que puede expresarse la autoridad y a ellas, *Wikipedia* les concede mucha importancia ya que son la garantía de que el usuario pueda acceder a la información utilizando su propia lengua y alfabeto. Las autoridades paralelas se registran en el código fuente de cada artículo de la enciclopedia y son mostradas en forma de hipervínculo dentro de un apartado específico del mismo. La figura 2 muestra el artículo de *Wikipedia* correspondiente a **Erasmus de Rotterdam**. Éste aparece encabezado por la autoridad en español (arriba), mientras que el acceso a las autoridades paralelas se hace desde el casillero denominado “en otros idiomas”, en el margen inferior izquierdo.

Este procedimiento permite que cada usuario acceda a la información buscando conforme a sus propios usos lingüísticos y alfabéticos. Esto es importante ya que, frente al modelo de control de autoridades de

los catálogos bibliográficos, *Wikipedia* adopta el modelo del control de acceso. El control de autoridades tradicional trata de determinar una única forma internacional autorizada para cada nombre. El control de acceso, sin embargo, admite como autoridades todas las posibles variantes idiomáticas que designan un mismo nombre, título o materia, y las vincula entre sí. Todas esas variantes se encuentran al mismo nivel, es decir, no hay una forma que prevalezca sobre otra por lo que es el usuario quien decide, según su conveniencia, la “puerta” por la que quiere acceder a la información¹¹.

En la práctica eso se traduce en que un usuario español que busque datos sobre **Erasmus de Rotterdam**, accedería a la información tecleando en el propio buscador de la enciclopedia o en cualquier buscador general *Erasmus de Rotterdam*, que es la forma más conocida para ese nombre en el ámbito hispano-parlante; un usuario francés, sin embargo, lo haría tecleando *Érasme*; uno alemán, *Erasmus von Rotterdam*; y uno ruso, *Эразм Роттердамский*. Desde un punto de vista diferente, un usuario de habla inglesa que buscara información en *Wikipedia* sobre la obra *Cien años de soledad* podría buscar por *One hundred years of solitude*, que es la forma por la que se conoce a esta novela de **García Márquez** en el ámbito anglosajón; un italiano por *Cent'anni di solitudine*, y un árabe por *مائة ا*



Figura 2

ةلزعلا نم ا ع. Como vemos, en ningún momento se “incomoda” al usuario obligándole a conocer el título original de la obra para acceder a la información, lo mismo que a un usuario español no se le obliga a que busque por *Der Zauberberg* para acceder a una información sobre el clásico de **Mann** *La montaña mágica*.

Lo anterior es posible porque cada artículo de *Wikipedia* incluye la relación completa de autoridades plurilingües que le dan título en las distintas ediciones de la enciclopedia. De este modo se conectan todos los artículos que tratan sobre el mismo tema, pero escritos en diferente idioma. Así, por ejemplo, si editamos el artículo correspondiente a *Erasmus de Rotterdam*, encontraremos la siguiente lista de autoridades paralelas (cada una aparece precedida del código que identifica el idioma):

```
[[als:Erasmus von Rotterdam]]
[[ar:إيراسموس روتردام]]
[[be-x-old:Эразм Ратэрдамскі]]
[[bg:Еразъм Ротердамски]]
[[bs:Erazmo Roterdamski]]
[[ca:Erasme de Rotterdam]]
[[cs:Erasmus Rotterdamský]]
[[cy:Desiderius Erasmus]]
[[da:Erasmus af Rotterdam]]
[[de:Erasmus von Rotterdam]]
[[el:Ντεζιντερίους Έρασμος]]
[[en:Desiderius Erasmus]]
[[eo:Erasmus de Roterdamo]]
[[et:Erasmus Rotterdamist]]
[[fa:إيراسموس روتردامی]]
[[fi:Erasmus Rotterdamlainen]]
[[fr:Érasme]]
[[fy:Erasmus]]
[[gl:Erasmus de Róterdam]]
[[he:סרטרם סומטרא]]
[[hr:Erazmo Roterdamski]]
[[hu:Rotterdami Erasmus]]
[[id:Desiderius Erasmus]]
[[io:Erasmus de Rotterdam]]
[[is:Desiderius Erasmus]]
[[it:Erasmus da Rotterdam]]
...
```

Como vemos por los dos ejemplos anteriores, el concepto de autoridad en *Wikipedia* es móvil y múltiple. Responde en todo momento a la forma con que el nombre es conocido en cada lengua. No existe una autoridad única admitida y con validez internacional. Es un modelo flexible, cómodo y directo. No podemos decir

lo mismo del modelo de los catálogos bibliográficos, ya que los estándares de catalogación estipulan formas muy rígidas y, a veces, no amparadas por el uso común. Baste decir que para el caso de **Erasmus**, la autoridad aceptada en un catálogo bibliográfico es siempre la forma latina del nombre, es decir, *Erasmus*, *Desiderius*, mientras que para el caso de la obra de **García Márquez**, la forma autorizada del título es siempre y para cualquier idioma *Cien años de soledad*. Convendremos que, comparativamente, los catálogos bibliográficos no salen muy bien parados en cuanto a su valoración como instrumentos pensados para el usuario.

El concepto de control de acceso no es exclusivo de *Wikipedia*, ni siquiera nuevo, pues fue formulado hace ya un par de décadas en el ámbito bibliotecario. Sin embargo, más allá de los debates de la *IFLA* sobre ficheros virtuales internacionales de autoridades¹², conocemos muy pocas aplicaciones prácticas, siendo *Wikipedia* una de las primeras. Un ejemplo gráfico de la diferencia entre control de autoridades (catálogos bibliográficos) y control de acceso (*Wikipedia*) se muestra en la figura 3.

2.3. Las *hatnotes*

Son notas o avisos breves que se colocan en la cabecera del artículo, antes del cuerpo del texto, y que sirven para: 1) desambiguar la autoridad de otras autoridades homónimas o conceptualmente muy relacionadas; o 2) definir la autoridad y explicar el alcance de la misma. Las *hatnotes* (notas sombrero) tienen la misma función que las notas de información de los registros de autoridad en los catálogos bibliográficos. Se introducen en cursiva y para su redacción se utilizan fórmulas predeterminadas que se recogen en los manuales de estilo de *Wikipedia*¹³. En la figura 4 se recogen algunos ejemplos.

2.4. Redirecciones

Juegan el mismo papel que las referencias de *véase* en un vocabulario controlado. Conducen al usuario desde una entrada no autorizada a una autorizada. Como en cualquier sistema de control de autoridades se crea una redirección por cada entrada no autorizada. Su finalidad contribuye a lograr el primer gran objetivo de cualquier sistema de autoridades: la agrupación o unificación de términos distintos que hacen referencia al mismo concepto.

En *Wikipedia* la redirección funciona de manera transparente. Si el usuario busca por una forma no autorizada, el sistema se encarga de llevarle a la página con la forma autorizada sin pasos intermedios o requerimientos de acciones por su parte. El único *feedback* que se recibe es un aviso (que se muestra debajo de la autoridad), donde se informa de que ha sido redirigido desde una variante que no es autorizada. Por ejemplo, si

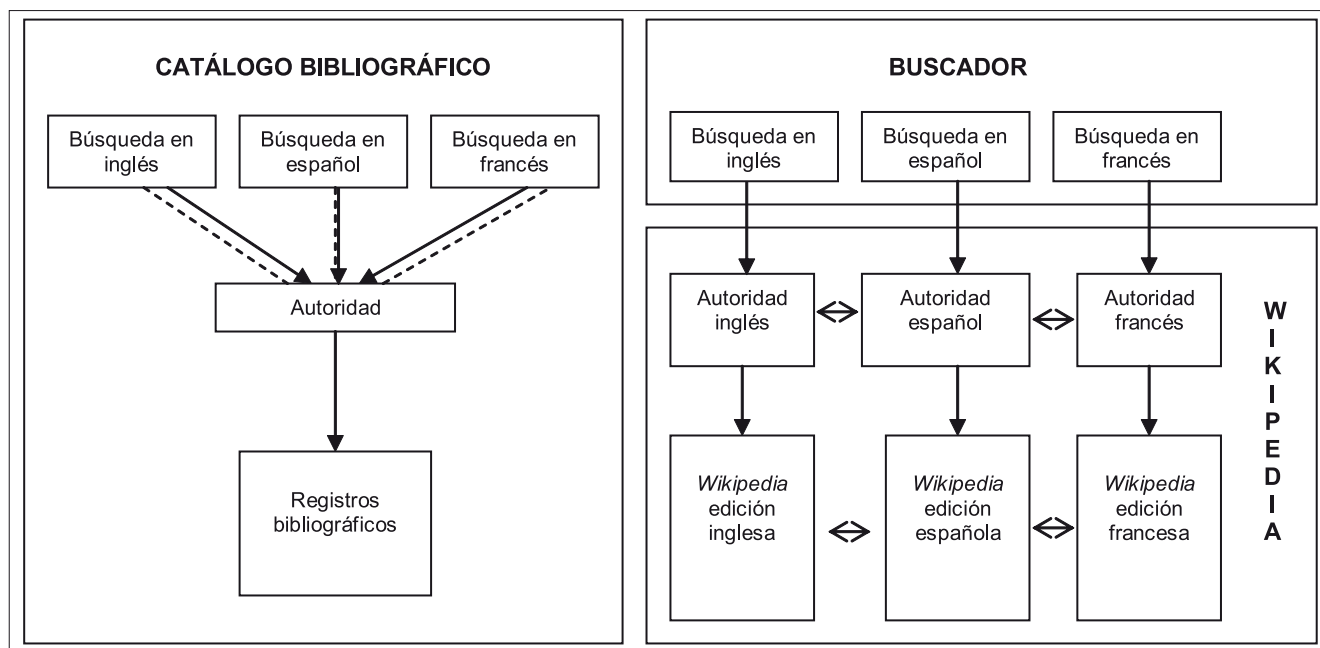


Figura 3

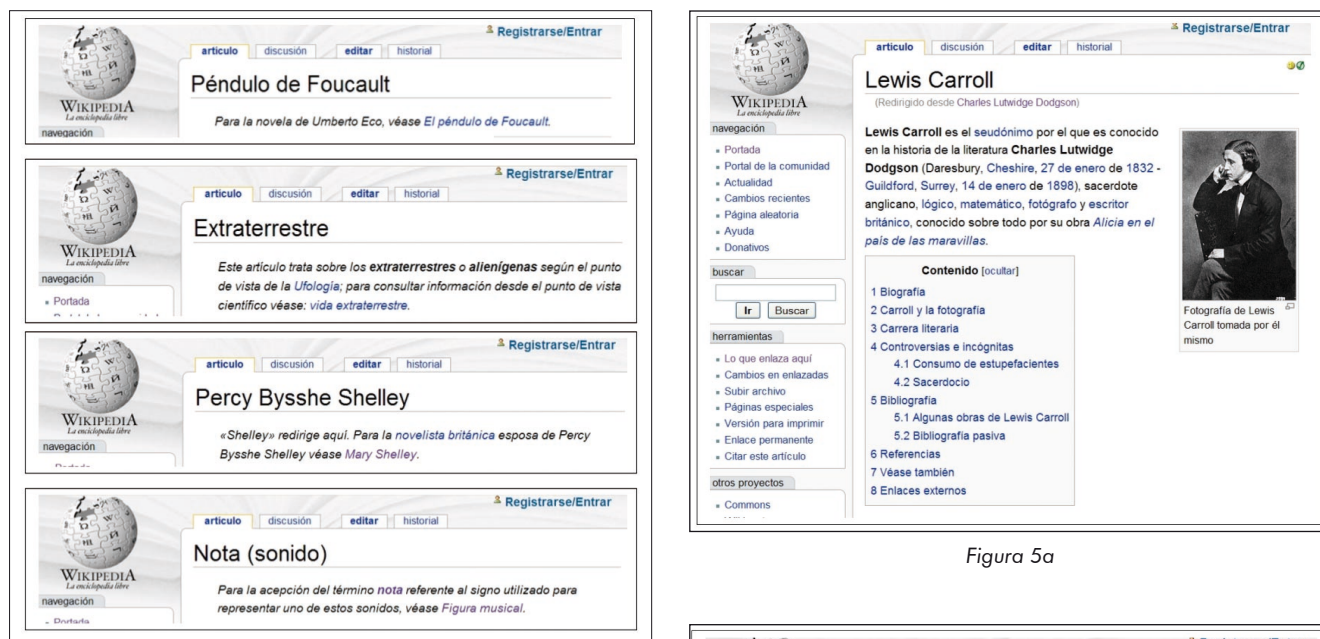


Figura 5a

Figura 4

buscamos **Charles Lutwidge Dodgson**, Wikipedia nos lleva a la página de **Lewis Carroll**, avisándonos de que hemos sido redirigidos desde la forma buscada (figura 5a). Esto es posible gracias a que existe una página de redirección que traza esa ruta (figura 5b).

Las variantes para las que Wikipedia crea una redirección están estipuladas en su manual *Redirects*¹⁴ y pueden ser tanto de nombre, seudónimos, nombres reales, nombres con título de nobleza, nombres religiosos, variantes ortográficas, de signos diacríticos, acrónimos, etc. Está estipulado que las variantes no autorizadas deben incluirse y mostrarse en las primeras líneas de texto del artículo (figura 6a). No obstante, el registro



Figura 5b

completo de variantes redirigidas a una autoridad puede verse con mayor fiabilidad en el menú herramientas, apartado "lo que enlaza aquí", opción "mostrar redirecciones" (figuras 6a y 6b).

2.5. Páginas de desambiguación

Son páginas de redirección múltiple, es decir, conducen al usuario desde una entrada no autorizada a más de una entrada autorizada. Se utilizan para deshacer homonimias y polisemias, el principal caballo de batalla



Figura 6a



Figura 6b

de cualquier vocabulario controlado. Cumplen por lo tanto el segundo gran objetivo de un sistema de autoridades: la diferenciación de términos que se escriben igual pero que tienen significados distintos.

Las páginas de desambiguación comienzan indicando que la entrada buscada es ambigua y que tiene varias acepciones. A continuación introducen las distintas acepciones mediante una breve frase explicativa que incluye un enlace a cada autoridad diferenciada.

Los wikipedistas disponen de un manual de procedimiento para la confección y actualización de páginas de desambiguación. En él se prescribe el uso de elementos para diferenciar autoridades (calificadores), el orden en que se han de mostrar, etc.¹⁵ *Wikipedia* da gran importancia a estas páginas-guía a juzgar por su número y su nivel de detalle, lo que dice mucho a favor de la calidad de su sistema de control de autoridades. En la figura 7 se muestra la página de desambiguación de *Wikipedia* para el término *mercurio*. Las autoridades no aparecen diferenciadas en esa página, sólo sus diferentes acepciones. Para ver la autoridad de cada artículo pulsaríamos el enlace que hay sobre cada término homónimo.

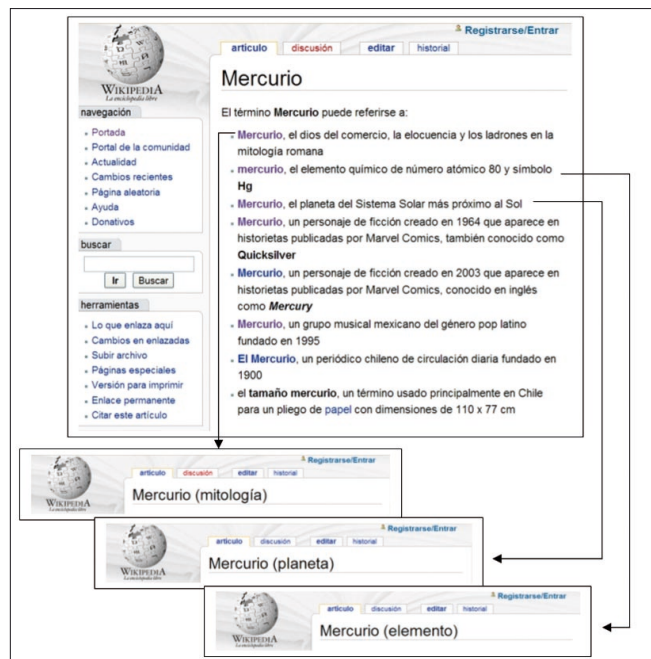


Figura 7

Cuando las acepciones que se presentan en la página de desambiguación son muchas, se presentan agrupadas por temas, para facilitar su consulta (figura 8).

Hay páginas de desambiguación en *Wikipedia* que no existen en los catálogos bibliográficos. Son verdaderas agrupaciones-diferenciaciones de nombres que tienen elementos de información en común. Son parecidas al resultado de un ojeo o *browsing* de autoridades como el que se puede hacer en cualquier catálogo bibliográfico, pero predeterminadas y con más información. No se trata de meros listados alfabéticos de autoridades, sino páginas que se van alimentando cada vez que un homónimo o cuasi-homónimo es añadido por un wikipedista a la enciclopedia. Véanse algunos ejemplos de estas páginas en la figura 9, donde se detallan



Figura 8



Figura 9

todas las autoridades que se llaman **Antonio García**, diferenciadas con fechas y una breve frase indicativa de su profesión o actividad; y en la figura 10, donde todos los **Cayo Julio César** son diferenciados con mínimos pero precisos datos de su genealogía. Es en estos casos donde se puede ver realmente la enorme utilidad y el potente alcance de estos elementos de conexión y organización de información.

2.6. Véase también

Mediante este recurso se consigue crear una verdadera tela de araña conectando autoridades cuyos significados están próximos, relacionados o pueden asociarse mentalmente por el usuario. El *véase también* cumple en cierto modo la función de las referencias de *véase además* de los catálogos bibliográficos, pero *Wikipedia* lo dota de una mayor funcionalidad. Si en los catálogos bibliográficos este procedimiento sólo se usa para re-

lacionar un mismo tipo de autoridades (materias entre sí, entidades entre sí, etc.), aquí es multidireccional, es decir, lo mismo relaciona un nombre de persona con una materia, una materia con una entidad, un nombre de lugar con un título, etc. De esta forma se consigue un consistente entramado conceptual, que nada tiene que ver con la modesta aspiración de los catálogos bibliográficos. En la figura 11 se muestra el área de *véase también* del artículo sobre **Albert Einstein**.



Figura 11

2.7. Categorías

Wikipedia define "categoría" como una agrupación de páginas que comparten algún tema en común y equipara su función a la que tendría un directorio en un sistema de archivos de un ordenador¹⁶. Las categorías facilitan la búsqueda de información, permitiendo ver el contenido de las páginas, o bien navegar a través de ellas gracias a su ordenación jerárquica. Se da tanta importancia a esta ubicación jerárquica que una autoridad/página no categorizada es considerada por *Wikipedia* como huérfana.

La categorización hace algo parecido a lo que las clasificaciones bibliográficas tradicionales (tipo *CDU*, *Dewey*, etc.) han venido haciendo desde hace siglos con los documentos, es decir, ubicarlos en clases temáticas. No obstante, existe una diferencia fundamental e insalvable entre la categorización y la clasificación. La clase se expresa siempre en lenguaje simbólico (notación) mientras que la categoría lo hace en lenguaje natural. Esto acerca a la categorización más al modo de proceder de un tesoro que al de una clasificación, en cuanto que organiza jerárquicamente los términos con la flexibilidad y eficacia que presentan los descriptores y sus relaciones de términos genéricos y específicos.

El concepto de jerarquía es entendido en la documentación tradicional como algo específicamente ligado a las materias. Muy rara vez la relación genérico-específico se ha trasladado al ámbito de las autoridades de



Figura 10

nombres de persona, de entidad o de título. *Wikipedia* aplica sin embargo la categorización a todas sus autoridades, cualquiera que sea su tipología. Ofrece así la posibilidad de tener agrupaciones de autoridades por categorías muy variadas. Algunos ejemplos, que resultarían insólitos en los catálogos bibliográficos, son: *Obras de Miguel Delibes*, *Reyes de España*, *Papas*, *Directores de cine estadounidenses*, *Músicos franceses*, *Entidades bancarias españolas*, *Cantantes de la década de 1970*, *Nacidos en 1950*, *Siglo XVIII*, etc. Esta variedad de campos semánticos convierte a la categorización de *Wikipedia* en una verdadera folksonomía, pero con el añadido positivo de la estructuración jerárquica. En la figura 12 vemos las categorías que se han asignado al artículo sobre la científica **Marie Curie**, y en la figura 13 se ofrece la presentación de la categoría *Deportes* con sus correspondientes subcategorías jerárquicas que nos permitirán (a través del símbolo +) descender en especificidad.



Figura 12



Figura 13

“Probablemente *Wikipedia* será un pilar de la web semántica, en detrimento de los modelos bibliotecarios de autoridades”

A modo de recopilación final, en el cuadro de la tabla 11 se recopilan las principales divergencias entre los modelos de control de autoridades de *Wikipedia* y los catálogos bibliográficos que han sido analizados en este trabajo.

CONTROL DE AUTORIDADES Principios divergentes	
Wikipedia	Catálogos bibliográficos
Normalización consensuada por los usuarios	Normalización impuesta por los bibliotecarios
Filosofía del control de acceso	Filosofía del control de Autoridades
Naturalidad	Artificialidad
Simplicidad de interpretación	Complejidad de interpretación
Flexibilidad en su aplicación	Rigidez en su aplicación
Autoridades en idioma del usuario	Autoridades en idioma original e idioma del usuario
Autoridades siempre en orden directo	Autoridades en orden directo/inverso
Calificación	Subordinación
No adición de términos al nombre	Adición de términos al nombre
Principio de autoridad única	Entrada múltiple
Preferencia del singular	Preferencia del plural

Tabla 11

3. Conclusiones

Wikipedia constituye una apuesta sólida por la calidad de la información y su organización en la Web. Tanto es así que ha desarrollado su propio sistema de control de autoridades para proveer a sus usuarios de un acceso cómodo y consistente al contenido de sus artículos. El análisis que hemos hecho en este artículo demuestra que se puede lograr una altísima eficacia en el control de autoridades con un modelo que se autogestiona, basado plenamente en las necesidades del usuario y regido por principios tales como la naturalidad, la simplicidad y la inexistencia de normas rígidas. El sentido común, la espontaneidad y la visión práctica del usuario definen este modelo novedoso y ponen en entredicho las supuestas bondades del control de autoridades tradicional, un modelo que, a ojos de hoy, se muestra esclerotizado por el peso de la tradición y por una escasa o nula adaptación a la tecnología.

Wikipedia está creciendo continuamente y aún le queda mucho camino por recorrer pero es de esperar que las soluciones de normalización que ha venido adoptando hasta ahora se vayan consolidando y perfeccionando. Las propuestas de futuro actuales para

internet pasan por el uso de ontologías para hacer una Red más inteligente que conduzca al procesamiento automático y a lo que se ha denominado web semántica. Por su tamaño, su potencia como conector de información y su alcance temático, *Wikipedia* se perfila, en serio detrimento de los modelos bibliotecarios de ficheros virtuales de autoridades, como un firme candidato a constituirse en uno de los pilares básicos de esa web semántica.

4. Referencias y notas

1. Es necesario señalar que otras enciclopedias con acceso web ofrecen cierto grado de normalización de su vocabulario, pero en ningún momento puede considerarse que hayan desarrollado un sistema de control de autoridades riguroso ni con el nivel de elaboración, alcance y extensión con el que se presenta hoy al usuario la enciclopedia *Wikipedia*. La tecnología web ha supuesto un importante replanteamiento de los procedimientos seguidos para la elaboración de las fuentes de información en general y de las enciclopedias en particular, sobre todo en lo que se refiere a la forma de presentar sus contenidos y de establecer y redactar sus entradas, pero de momento sólo *Wikipedia* ofrece un coherente y sofisticado control de vocabulario capaz de competir con el tradicional control de autoridades bibliotecario y, por ello, susceptible de ser estudiado desde una óptica profesional, tal como pretendemos hacer en el presente artículo.

2. *Wikipedia, la enciclopedia libre*
<http://es.wikipedia.org/wiki/Wikipedia>

3. El principal manual para autoridades de *Wikipedia* es *Wikipedia: naming conventions*. Aunque éste sirve de modelo para las demás ediciones de la enciclopedia, la edición española tiene sus propias convenciones recogidas en el manual *Wikipedia: convenciones de títulos*.
http://en.wikipedia.org/wiki/Wikipedia:Naming_conventions
http://es.wikipedia.org/wiki/Wikipedia:Convenciones_de_t%C3%ADtulos

4. En este contexto, el término bibliotecario puede resultar ambiguo, ya que en el ámbito español de *Wikipedia* se denomina bibliotecario a la persona que cuenta con permiso para administrar contenidos de la enciclopedia. Aquí obviamente nos referimos al profesional de la biblioteconomía.

5. *Wikipedia: naming conventions*
http://en.wikipedia.org/wiki/Wikipedia:Naming_conventions

6. *Wikipedia: naming conventions (people)*
[http://en.wikipedia.org/wiki/Wikipedia:Naming_conventions_\(people\)](http://en.wikipedia.org/wiki/Wikipedia:Naming_conventions_(people))

7. *Reglas de catalogación*. Ed. nuevamente rev., 5ª reimp. Madrid: Ministerio de Cultura, 2006. La cifra que damos resulta de sumar todas las reglas, subreglas y especificaciones que recoge en sus capítulos 15, 16 y 17.

8. Por poner un ejemplo, la normativa para la redacción de los encabezamientos de materia de la *Biblioteca Nacional de España* recoge más de 150 reglas y especificaciones distintas. Véase *Encabezamientos de materia: normativa para su redacción*. Madrid: Ministerio de Cultura, 1991.

9. Todos los ejemplos de entradas de *Wikipedia* que se muestran en este trabajo han sido tomados de su edición española. Por su parte, los ejemplos de puntos de acceso normalizados de catálogos bibliográficos han sido obtenidos de los catálogos de: *Biblioteca Nacional de España*, *Red de Bibliotecas Universitarias Españolas (Rebiun)* y *Red de Bibliotecas del Csic*.

10. En 1841, **Anthony Panizzi**, bibliotecario jefe del *British Museum*, dirigió un comité para redactar las reglas para la compilación del nuevo catálogo de la biblioteca del *Museo* (la que devino *British Library*). Las *Rules for the compilation of the catalogue of printed books*, conocidas como "las 91 reglas de Panizzi" son consideradas como el primer código moderno de catalogación. Sus reglas estaban diseñadas para la confección de un catálogo alfabético y sistematizaban por primera vez el orden de entrada de las autoridades. En: **García Arilla, M. R.** *Teoría e historia de la catalogación de documentos*. Madrid: Síntesis, 1996.

11. Para un estudio más amplio de las nuevas tendencias en el control de autoridades en el siglo XXI, concretamente sobre el concepto de control de acceso, véase **Jiménez Pelayo, J.** y **García Blanco, R.** *El catálogo de autoridades: creación y gestión en unidades documentales*. Gijón: Síntesis, 2002, p. 66 y ss.

12. En el ámbito bibliotecario, la *Deutsche Bibliothek*, la *Library of Congress* y *OCLC* están desarrollando conjuntamente un Fichero de Autoridades Virtual Internacional (*VIAF*) de nombres de persona, que relaciona registros de autoridad de las agencias bibliográficas nacionales del mundo, con objeto de que esté libremente accesible en internet. Su meta se fija a largo plazo. Véase **Bennet, R.** et al. "VIAF (Fichero de Autoridades Virtual Internacional)". En: *72nd IFLA General conference and council, 20-24 August 2006, Seoul, Korea*.
http://www.ifla.org/IV/ifla72/papers/123-Bennett_trans-es.pdf

13. *Wikipedia: hatnote*
<http://en.wikipedia.org/wiki/Wikipedia:Hatnotes>

14. *Wikipedia: redirecciones*
<http://es.wikipedia.org/wiki/Wikipedia:Redirecciones>

15. *Wikipedia: página de desambiguación*
http://es.wikipedia.org/wiki/Wikipedia:P%C3%A1gina_de_desambiguaci%C3%B3n

16. *Wikipedia: categoría*
<http://es.wikipedia.org/wiki/Ayuda:Categor%C3%ADa>

Jesús Jiménez-Pelayo, Biblioteca Virtual de Andalucía, Granada.
jesusj.jimenez@juntadeandalucia.es

Quieres recibir tu copia del Anuario ThinkEPI 2008 en casa*?

ANUARIO

Think
EPI

Haz tu pedido en <http://www.thinkepi.net/anuario.html>

¡Sólo 35€!

*La tarifa individual particular sólo sufraga el coste marginal o incremental de la imprenta. Está pensada para los particulares que se costean el Anuario de su bolsillo, bien porque quieren leerlo en su casa o bien porque su institución no lo ha comprado.